

Variable-Length Feedback Codes under a Strict Delay Constraint

Seong Hwan Kim, *Member, IEEE*, Dan Keun Sung, *Fellow, IEEE*, and Tho Le-Ngoc, *Fellow, IEEE*

Abstract—We study variable-length feedback (VLF) codes under a strict delay constraint to maximize their average transmission rate (ATR) in a discrete memoryless channel (DMC) while considering periodic decoding attempts. We first derive a lower bound on the maximum achievable ATR, and confirm that the VLF code can outperform non-feedback codes with a larger delay constraint. We show that for a given decoding period, as the strict delay constraint, L , increases, the gap between the ATR of the VLF code and the DMC capacity scales at most on the order of $O(L^{-1})$ instead of $O(L^{-1/2})$ for non-feedback codes as shown in Polyanskiy *et al.* [“Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307-2359, May 2010.]. We also develop an approximation indicating that, for a given L , the achievable ATR increases as the decoding period decreases.

Index Terms—VLF codes, strict delay constraint, DMC.

I. INTRODUCTION

THE channel capacity, i.e., the maximum coding rate with an arbitrarily small error probability, needs an assumption of infinite block-length [1]. However, the block error probability (BEP) is non-zero in practice since the block-length must be finite. For given block-length and BEP constraints, lower and upper bounds on the maximum coding rate were found in [2]. Meanwhile, feedback is known to be very useful for performance improvement if variable-length codes are applied. In the regime of asymptotically long average block-length, the error exponent, regarded as the exponential rate of decay of the BEP with respect to the average block-length, has been an important performance measure for variable-length codes with feedback [3]–[9]. However, the analysis cannot present the maximum coding rate for fixed BEP and fixed average block-length.

Recently, Polyanskiy *et al.* [10] focused on the regime of fixed BEP and fixed average block-length in their study of variable-length feedback (VLF) codes over discrete memoryless channels (DMC), formulating the following problem:

$$\max M \quad \text{s.t.} \quad \Pr[W \neq \widehat{W}] \leq \epsilon, \quad \mathbb{E}[\tau] \leq L, \quad (1)$$

where M , W , \widehat{W} , and τ denote the number of codewords, transmitted message, estimated message, and the length of a

received sequence. The positive real values ϵ and L denote constraints. A transmitter sends a codeword as time-domain symbols and a receiver attempts to decode it at every symbol reception and sends a stop-feedback to the transmitter when the receiver decides to decode it. Therefore, τ also represents a delay. They showed that a significant spectral efficiency gain can be obtained by using VLF codes for given average delay and BEP constraints as compared to the non-feedback fixed-length codes. They also considered a strict delay constraint in formulating the following problem:

$$\max M \quad \text{s.t.} \quad \Pr[\{W \neq \widehat{W}\} \cup \{\tau > L\}] = \epsilon. \quad (2)$$

Analyzing (2), they alleged that the strict delay constraint nullifies the gain of VLF codes over the non-feedback fixed-length codes. However, problem (2) does not maximize the average transmission rate (ATR), $\frac{\log_2 M}{\mathbb{E}[\tau]}$, which is also defined as the spectral efficiency. Therefore, it is not certain whether VLF codes have a gain or not over the non-feedback codes in terms of the spectral efficiency, under a strict delay constraint. Besides VLF codes, they also studied VLF with termination (VLFT) codes in which the transmitter decides to stop transmitting a message by observing the output of the decoder fed back from the receiver and sends a termination signal to the receiver under an assumption of error-free feedback and feed-forward channels. Chen *et al.* [11] added practical constraints to Polyanskiy’s VLFT codes by limiting the length of codewords and attempting to decode a codeword only at specified times. In their setting, the same codeword is reused if the receiver fails to decode it after receiving the whole original codeword while the average delay is limited as given in problem (1). However, the VLFT codes are far from practical because existing feedback codes such as Hybrid Automatic Repeat & reQuest (HARQ) schemes only use acknowledgement (ACK)/negative-ACK (NACK) feedbacks.

In this letter, we revisit VLF codes under a strict delay constraint and a fixed BEP to maximize the achievable ATR (instead of the number of codewords). The regime of strict (rather than *average*) delay constraints is desired to guarantee delay requirements of real-time traffic applications. We also consider periodic decoding attempts where the receiver attempts to decode a codeword periodically. We derive the achievable ATR as the lower bound on the maximum ATR of VLF codes. We prove that gap between the ATR of VLF codes and the DMC capacity scales at most on the order of $O(L^{-1})$ for a given decoding period instead of $O(L^{-1/2})$ for non-feedback codes with fixed length of L as shown in [2]. This comparison shows a significant gain of VLF codes over the non-feedback codes under a strict delay constraint. We also derive an approximation of the achievable ATR expression indicating that, for a given L , the achievable ATR increases as the decoding period decreases. In [12], we studied HARQ

Copyright ©2015 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

This manuscript has been accepted for publication in IEEE Communications Letters.

S. H. Kim and T. Le-Ngoc are with the Department of Electrical & Computer Engineering, McGill University, Canada (E-mail: seonghwan.kim@mcgill.ca; tho.le-ngoc@mcgill.ca). D. K. Sung is with the Department of EE, KAIST, Korea (dksung@ee.kaist.ac.kr).

Digital Object Identifier 10.1109/LCOMM.2015.2398866

schemes under a strict delay constraint in AWGN channels for a given BEP with the following two ideal assumptions: 1) the decoding period is sufficiently long; 2) the NACK event is identical to the error event. In this letter, without the above two assumptions, we find the theoretical bound of VLF codes in DMC.

Throughout this letter, we use the following notation: X , x , and P_X denote a random variable, its sample value, and the probability distribution of X , respectively. $x^n = (x_1, x_2, \dots, x_n)$ denotes an n -dimensional vector and x_j the j th element of x^n . $\log(x) = \log_e(x)$ unless otherwise stated.

II. PROBLEM STATEMENT

In this section, we briefly introduce the channel and VLF codes which are modified from those in [10] for defining periodic decoding attempts, and formulate our optimization problem.

Channel: A DMC consists of a pair of input X and output Y on the finite alphabets \mathcal{A} and \mathcal{B} , respectively, with a conditional probability, $P_{Y_i|X_i} = P_{Y_1|X_1}, \forall i \geq 1$.

(l, d) VLF code: An (l, d) VLF code with M messages, maximum allowable number of decoding attempts, l , and decoding period (i.e., interval between 2 consecutive decoding attempts), d , is defined as follows:

- 1) A random variable $U \in \mathcal{U}$ with a probability distribution of P_U represents a codebook shared by both transmitter and receiver.
- 2) A sequence of encoders $f_n : \mathcal{U} \times \{1, \dots, M\} \rightarrow \mathcal{A}$ represent the channel input at time n , $X_n = f_n(U, W)$ where $W \in \{1, \dots, M\}$ is the equi-probable message.
- 3) A sequence of decoders $g_k : \mathcal{U} \times \mathcal{B}^{dk} \rightarrow \{1, \dots, M\}$ attempting to provide the estimate of W at time dk where k denote the number of decoding attempts, respectively.
- 4) A final decision is made by the receiver at a stopping time $d\tau^* : \widehat{W} = g_{\tau^*}(U, Y^{d\tau^*})$.

Optimization problem: For an (l, d) VLF code with $0 < \epsilon < 1$, we maximize its achievable ATR under a strict delay constraint as follows:

$$\mathcal{T}_f^*(l, d, \epsilon) = \max_M \frac{\log M}{d\mathbb{E}[\tau^*]} \quad \text{s.t.} \quad \begin{cases} \Pr[W \neq \widehat{W}] \leq \epsilon, \\ \Pr[\tau^* \leq l] = 1. \end{cases}$$

$L = dl$ represents the limited length of the received sequence (i.e., the delay). Setting $d = 1$ indicates the case of attempting to decode a codeword at every symbol reception.

III. ACHIEVABILITY ANALYSIS

For achievability analysis, we specify the codebook, encoder, and decoder which are modified from those in [10] by adding functions to satisfy strict delay constraints as follows:

Codebook: A codebook U is defined on space \mathcal{U} such as

$$\mathcal{U} \triangleq \underbrace{\mathcal{A}^{dl} \times \dots \times \mathcal{A}^{dl}}_{M \text{ times}}, \quad P_U \triangleq \underbrace{P_X^{dl} \times \dots \times P_X^{dl}}_{M \text{ times}},$$

where X is distributed according to P_X on \mathcal{A} . The realization of U defines M dl -dimensional vectors $\mathbf{C}_j \in \mathcal{A}^{dl}, j = 1, \dots, M$.

Encoder: The encoding sequence f_n maps an equi-probable message j to $\mathbf{C}_j \in \mathcal{A}^{dl}$ and provides the channel input at time n ,

$$f_n(j) = (\mathbf{C}_j)_n \quad \text{for } 1 \leq n \leq dl,$$

where $(\mathbf{C}_j)_n$ is the n th coordinate of the vector \mathbf{C}_j .

Decoder: A decoder computes the j -th information density at the k -th decoding attempt for $1 \leq k \leq l$,

$$S_{j,k} \triangleq i(\mathbf{C}_j(dk), Y^{dk}), \quad j = 1, \dots, M,$$

where $\mathbf{C}_j(n)$ is the first n symbols of \mathbf{C}_j and, the information density between x^n and y^n is defined as

$$i(x^n; y^n) = \log \frac{dP_{Y^n|X^n}(y^n|x^n)}{dP_{Y^n}(y^n)}.$$

The decoder defines M stopping instances computing the information densities

$$\tau_j \triangleq \begin{cases} \inf\{0 \leq k \leq l : S_{j,k} \geq \gamma\} & \text{if } \bigvee_{k=1}^l \{S_{j,k} \geq \gamma\} \\ l+1 & \text{Otherwise,} \end{cases}$$

where γ is the threshold to stop computing information density for each codeword, and for statements A and B, $A \vee B$ (resp., $A \wedge B$) is true if A or B (resp., A and B) are true. Since $\bigwedge_{k=1}^l \{S_{j,k} < \gamma\}$ can be checked at time l , a mapping $\tau_j = l+1$ does not mean that the observation at the $(l+1)$ -th decoding attempt is required. Define a random variable

$$\tau' \triangleq \min\{\tau_1, \dots, \tau_M\}. \quad (3)$$

The final decision is made by the decoder at the stopping time

$$\tau^* \triangleq \begin{cases} \tau' & \text{if } \tau' \leq l \\ l & \text{if } \tau' = l+1, \end{cases} \quad (4)$$

i.e., τ^* is always smaller than or equal to l . The output of the decoder is

$$g(Y^{d\tau^*}) \triangleq \begin{cases} \max\{j : \tau_j = \tau'\} & \text{if } \tau' \leq l \\ M & \text{if } \tau' = l+1. \end{cases} \quad (5)$$

For the VLF code specified above, we find a lower bound on $\mathcal{T}_f^*(l, d, \epsilon)$.

Lemma 1: For an arbitrary DMC with capacity C , interval d , and $\exp\{-Cdl/2\} < \alpha < \epsilon$, the maximum M satisfying $\Pr[W \neq \widehat{W}] \leq \epsilon$ and $\tau^* \leq l$ is lower-bounded by

$$\max_{\substack{M \\ \Pr[W \neq \widehat{W}] \leq \epsilon \\ \tau^* \leq l}} M \geq \left\lceil (\epsilon - \alpha) \exp \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) Cdl \right\} + 1 \right\rceil.$$

Proof: Let us define the following instances:

$$\tau = \inf\{k \geq 1 : i(X^{dk}, Y^{dk}) \geq \gamma\}, \quad (6a)$$

$$\bar{\tau} = \inf\{k \geq 1 : i(\bar{X}^{dk}, Y^{dk}) \geq \gamma\}, \quad (6b)$$

where X and \bar{X} are independent random variables with the same distribution and Y is the channel output when X

is channel input. The average probability of error with M codewords is upper-bounded as

$$\mathbb{P}[g(Y^{d\tau^*}) \neq W] \leq \mathbb{P}[g(Y^{d\tau^*}) \neq 1 | W = 1] \quad (7)$$

$$= \mathbb{P} \left[\bigcup_{j=2}^M \{\tau_j \leq \tau_1 \leq l\} \bigcup \{l < \tau_1\} \middle| W = 1 \right] \quad (8)$$

$$\leq (M-1)\mathbb{P}[\tau_2 \leq \tau_1 \leq l | W = 1] + \mathbb{P}[l < \tau_1 | W = 1] \quad (9)$$

$$= (M-1)\mathbb{P}[\bar{\tau} \leq \tau \leq l] + \mathbb{P}[l < \tau] \quad (10)$$

$$\leq (M-1)\mathbb{P}[\bar{\tau} \leq \tau] + \mathbb{P}[l < \tau], \quad (11)$$

where (7) follows from (5), (8) from (3), (4), and (5), (9) from a union bound and symmetry for τ_j , (10) from (6).

Note that, as compared to (45) in [10], (11) contains the additional term $\mathbb{P}[l < \tau]$ for the event that no codeword is detected within l decoding attempts. $\mathbb{P}[\bar{\tau} \leq \tau]$ is upper-bounded by (see (111)-(118) of [10]¹)

$$\mathbb{P}[\bar{\tau} \leq \tau] \leq \exp\{-\gamma\}. \quad (12)$$

$\mathbb{P}[l < \tau]$ is upper-bounded by

$$\mathbb{P}[l < \tau] = \mathbb{P} \left[\bigcap_{k=1}^l \{i(X^{dk}, Y^{dk}) < \gamma\} \right] \quad (13)$$

$$\leq \mathbb{P}[i(X^{dl}, Y^{dl}) < \gamma] \quad (14)$$

$$= \mathbb{P}[i(X^{dl}, Y^{dl}) < (1-\delta)Cdl] \quad (15)$$

$$\leq \exp\{-\delta^2 Cdl/2\},$$

where (13) follows from (6a), (14) is obtained by setting $\gamma = (1-\delta)Cdl$ for $C = \mathbb{E}[i(X; Y)]$ and $0 < \delta < 1$, and (15) is obtained by using a Chernoff bound.

Instead of $\mathbb{P}[g(Y^{d\tau^*}) \neq W] \leq \epsilon$, we set a stricter constraint using (11) and substitute (12) and (15) into (11),

$$(M-1)\exp\{-(1-\delta)Cdl\} + \exp\{-\delta^2 Cdl/2\} \leq \epsilon. \quad (16)$$

Notice that any M satisfying (16) also satisfies $\mathbb{P}[g(Y^{d\tau^*}) \neq W] \leq \epsilon$. To satisfy (16), let $\exp\{-\delta^2 Cdl/2\} = \alpha$ for $0 < \alpha < \epsilon$. For $0 < \delta < 1$, the range of α becomes $\exp\{-Cdl/2\} < \alpha < \epsilon$. Therefore, (16) can be rewritten as

$$M \leq (\epsilon - \alpha) \exp \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) Cdl \right\} + 1$$

for $\exp\{-Cdl/2\} < \alpha < \epsilon$. We know that $\hat{M}(\alpha) = \lfloor (\epsilon - \alpha) \exp\{(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}})Cdl\} + 1 \rfloor$ is the maximum M satisfying (16). In other words, we can achieve $\hat{M}(\alpha)$ with $\Pr[W \neq \hat{W}] \leq \epsilon$ and $\Pr[\tau^* \leq l]$. ■

Lemma 2: For an arbitrary DMC with capacity C , interval d , and $\exp\{-Cdl/2\} < \alpha < \epsilon$, $\mathbb{E}[\tau^*]$ is upper-bounded as

$$\mathbb{E}[\tau^*] \leq \min \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) l + \frac{a_0}{C}, l \right\},$$

where a_0 is the maximum value of $i(X, Y)$.

¹Polyanskiy et al. proved (12) only for $d = 1$. However, the proof is easily extended to the cases of $d > 1$.

Proof:

$$\mathbb{E}[\tau^*] \leq \min(\mathbb{E}[\tau'], l) \quad (17)$$

$$\leq \min \left(\frac{1}{M} \sum_{j=1}^M \mathbb{E}[\tau_j | W = j], l \right) \quad (18)$$

$$= \min(\mathbb{E}[\tau_1 | W = 1], l) \quad (19)$$

$$= \min(\mathbb{E}[\tau], l), \quad (20)$$

where (17) follows from (4), (18) follows from (3), (19) follows from symmetry, and (20) follows from (6a). Since $i(X^{dk}, Y^{dk}) - dkI(X, Y) = i(X^{dk}, Y^{dk}) - dkC$ is a martingale, from Doob's optional stopping theorem [10], we obtain $\mathbb{E}[i(X^{d\tau}, Y^{d\tau}) - d\tau C] = 0$. Then, $\mathbb{E}[\tau]$ is upper-bounded as

$$\mathbb{E}[\tau] = \frac{\mathbb{E}[i(X^{d\tau}, Y^{d\tau})]}{dC} \leq \frac{\gamma(\alpha) + da_0}{dC}, \quad (21)$$

where (21) follows from (6a) and $\gamma(\alpha) = \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) Cdl$. Note that (20) is similar to (39) of [10] except for the limitation to l and (21) is similar to (107) of [10] except for d and $\gamma(\alpha)$. ■

Theorem 1: For an arbitrary DMC with capacity C , and $e^{-\frac{Cdl}{2}} < \epsilon < 1$,

$$\mathcal{T}_f^*(l, d, \epsilon) \geq \max_{e^{-\frac{Cdl}{2}} < \alpha < \epsilon} \frac{\log(\epsilon - \alpha) + Cdl - \sqrt{2Cdl \log(1/\alpha)}}{d \min \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) l + \frac{a_0}{C}, l \right\}}. \quad (22)$$

Proof: From Lemmas 1 and 2,

$$\begin{aligned} \max_M \frac{\log M}{d\mathbb{E}[\tau^* | M]} &\geq \max_{e^{-\frac{Cdl}{2}} < \alpha < \epsilon} \frac{\log \hat{M}(\alpha)}{d\mathbb{E}[\tau^* | \hat{M}(\alpha)]} \\ &\geq \max_{e^{-\frac{Cdl}{2}} < \alpha < \epsilon} \frac{\log \left[(\epsilon - \alpha) \exp \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) Cdl \right\} + 1 \right]}{d \min \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) l + \frac{a_0}{C}, l \right\}} \\ &\geq \max_{e^{-\frac{Cdl}{2}} < \alpha < \epsilon} \frac{\log(\epsilon - \alpha) + Cdl - \sqrt{2Cdl \log(1/\alpha)}}{d \min \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}} \right) l + \frac{a_0}{C}, l \right\}}. \end{aligned}$$

While Theorem 1 offers the expression to compute the lower bound on the maximum ATR in terms of C, l, d , and ϵ , the expression does not provide a simple view of the effect of l on $\mathcal{T}_f(l, d, \epsilon)$. The following theorem establishes the asymptotic behavior of the gap between C and $\mathcal{T}_f^*(l, d, \epsilon)$ as l increases.

Theorem 2: For an arbitrary DMC with capacity C and any $e^{-\frac{Cdl}{2}} < \epsilon < 1$, the gap between the capacity and the maximum ATR of a VLF code scales at most,

$$\Delta_{\text{VLF}} \triangleq C - \mathcal{T}_f^*(l, d, \epsilon) = O(1/l),$$

where $f(n) = O(g(n))$ iff there are constants c and n_0 such that $f(n) \leq cg(n) \forall n > n_0$.

Proof: $\Delta_{\text{VLF}} = C - \max_M \frac{\log M}{d\mathbb{E}[\tau^*|M]}$ is upper-bounded as

$$\Delta_{\text{VLF}} \leq \min_{e^{-\frac{Cdl}{2}} < \alpha < \epsilon} C - \frac{\log(\epsilon - \alpha) + Cdl - \sqrt{2Cdl \log(1/\alpha)}}{d \min \left\{ \left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}}\right) l + \frac{a_0}{C}, l \right\}} \quad (23)$$

$$\leq \min_{e^{-\frac{Cdl}{2}} < \alpha < \min\{\epsilon, e^{-\frac{da_0^2}{2Cl}}\}} \frac{da_0 - \log(\epsilon - \alpha)}{\left(1 - \sqrt{\frac{2 \log(1/\alpha)}{Cdl}}\right) dl + \frac{da_0}{C}}, \quad (24)$$

where (23) follows from Theorem 1 and (24) follows from limiting the range of α by eliminating min in the denominator and from the relationship, $\beta_0 - \beta_1/\beta_2 = (\beta_2 - \beta_1/\beta_0)\beta_0/\beta_2$. For terms related to l on the right-hand-side (RHS) of (24), we obtain the following three convergences:

$$\lim_{l \rightarrow \infty} \exp\{-Cdl/2\} = 0, \quad \lim_{l \rightarrow \infty} \exp\{-a_0^2 d/(2Cl)\} = 1, \\ \lim_{l \rightarrow \infty} \frac{dl - \sqrt{2dl \log(1/\alpha)/C} + da_0/C}{l} = d.$$

Therefore, there exist constants l_1 and l_2 such that $\exp\{-\frac{Cdl}{2}\} < \frac{\epsilon}{2}$ for any $l > l_1$ and $\frac{\epsilon}{2} < \exp\{-\frac{a_0^2 d}{2Cl}\}$ for any $l > l_2$, respectively. In addition, by setting $\alpha = \epsilon/2$, for a given constant $0 < c_0 < 1$, there exists a constant $l_3 > \max\{l_1, l_2\}$ such that $\frac{dl - \sqrt{2dl \log(2/\epsilon)} + da_0/C}{l} > c_0 d$ for any $l > l_3$. Finally, we have for $l > l_3$

$$\Delta_{\text{VLF}} \leq \frac{da_0 - \log(0.5\epsilon)}{dl - \sqrt{2dl \log(2/\epsilon)} + \frac{da_0}{C}} \leq \frac{da_0 - \log(0.5\epsilon)}{ldc_0}. \quad (25)$$

Therefore, $\Delta_{\text{VLF}} = O(1/l)$. ■

Remarks:

- 1) For given d , since L is linearly proportional to l , Δ_{VLF} scales as $\Delta_{\text{VLF}} = O(1/L)$ from Theorem 2.
- 2) From (25), the achievable ATR expression is approximated for $L \gg d$ as

$$\mathcal{T}_f^*(l, d, \epsilon) \approx C - \frac{da_0 - \log(0.5\epsilon)}{L}. \quad (26)$$

- 3) From (26), for a given L , the approximation of $\mathcal{T}_f^*(l, d, \epsilon)$ increases as d decreases.

IV. ILLUSTRATIVE RESULTS

We show some illustrative results obtained from the analysis in Section III in terms of ATR [b/s/Hz]. To use b/s/Hz as a unit of measure in illustrative results, quantities related with the spectral efficiency are divided by $\log_e 2$.

Fig. 1 plots the ATR [b/s/Hz] versus L for binary symmetric channel (BSC) with crossover probability $q = 0.11$, $\epsilon = 10^{-3}$, and $C = 0.5$ b/s/Hz for $d = 1, 50$, and 100 . In BSC(q) channel, $C = \log(2) - H(q)$ and $a_0 = \max\{\log(2q), \log(2(1-q))\}$. The legend ‘No feedback’ corresponds to the fixed blocklength codes without feedback given by Theorem 52 of [2]. It is interesting to observe that the ATR of VLF codes converges to the capacity faster than that of the fixed-length codes even under a strict delay constraint. When $d = 1$, to achieve 90% of the capacity, $L = 360$ is enough for VLF-code while at least 3100 is required for non-feedback code. The ATRs

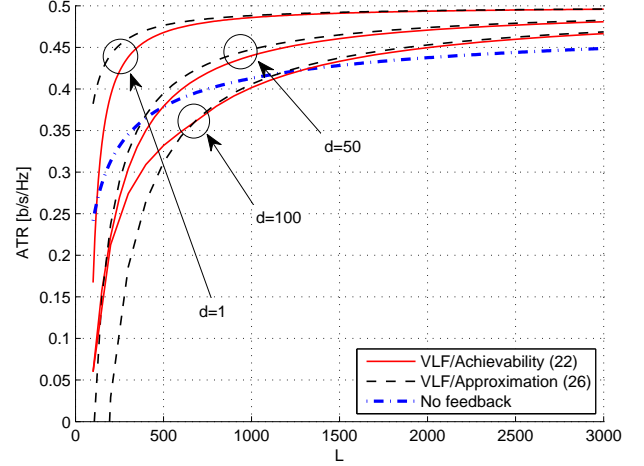


Fig. 1. The ATR vs. L for the BSC(0.11); probability of error $\epsilon = 10^{-3}$ for $d = 1, 50$, and 100 .

of the VLF code are lower than those of the non-feedback code for $L \leq 120$ with $d = 1$, $L \leq 450$ with $d = 50$, and $L \leq 1200$ with $d = 100$. In VLF codes, decoding attempts at $k < l$ give an opportunity to stop the transmission before l . However, these attempts simultaneously increase the probability that other messages than the transmitted one are detected as an estimated message. As L increases, the positive effect of VLF codes dominates the negative effect. Therefore, as L increases, VLF codes offer much higher ATR than non-feedback codes. We can also observe that, as d decreases, the ATR increases for a given L as indicated by (26). In practice, smaller d may need faster processing or incur higher complexity, which should be considered in the design of VLF codes for a specific application. As shown in Fig. 1, the approximation (26) provides results very close to those of (22).

V. CONCLUSIONS

We investigated a VLF code under a *strict delay constraint* to maximize the ATR in DMCs while considering periodic decoding attempts. We first derived a lower bound on the maximum achievable ATR and showed that the VLF code can outperform the non-feedback codes for increasing L . We also proved that the gap between the ATR of the VLF code and the DMC capacity scales at most on the order of $O(L^{-1})$ for a given d instead of $O(L^{-1/2})$ for non-feedback codes as previously shown in [2]. The approximated expression of the ATR indicates that the ATR increases with decreasing decoding period d .

REFERENCES

- [1] T. M. Cover and J. A. Thomas, “Elements of information theory, 2nd ed.” Wiley-Interscience, 2006.
- [2] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Channel coding rate in the finite blocklength regime,” *IEEE Trans. Inform. Theory*, vol. 56, no. 5, May 2010.
- [3] M. V. Burnashev, “Data transmission over a discrete channel with feedback. random transmission time,” *Probl. Inf. Transm.*, vol. 12, no. 4, pp. 10–30, 1976.

- [4] —, “Sequential discrimination of hypotheses with control of observations,” *Math. USSR, Izvestia*, vol. 15, no. 3, p. 419440, 1980.
- [5] H. Yamamoto and K. Itoh, “Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback,” *IEEE Trans. Inform. Theory*, vol. 25, no. 6, pp. 729–733, Nov. 1979.
- [6] N. Shulman, “Communication over an unknown channel via common broadcasting,” *Ph.D. dissertation, Tel-Aviv Univ., Tel-Aviv*, vol. 25, no. 6, pp. 729–733, Nov. 2003.
- [7] S. C. Draper, B. J. Frey, and F. R. Kschischang, “Efficient variable length channel coding for unknown DMCs,” in *Proc. IEEE ISIT*, 2004.
- [8] A. Tchamkerten and E. Telatar, “A feedback strategy for binary symmetric channels,” in *Proc. IEEE ISIT*, Jun. 2002.
- [9] —, “Optimal feedback schemes over unknown channels,” in *Proc. IEEE ISIT*, Jun. 2004.
- [10] Y. Polyanskiy, H. V. Poor, and S. Verdú, “Feedback in the non-asymptotic regime,” *IEEE Trans. Inform. Theory*, vol. 57, no. 8, pp. 4903–4925, Aug. 2011.
- [11] T. Chen, A. R. Williamson, and R. D. Wesel, “Variable-length coding with feedback: Finite-length codewords and periodic decoding,” in *Proc. IEEE ISIT*, July 2013.
- [12] S. H. Kim, D. K. Sung, and T. Le-Ngoc, “Performance analysis of incremental redundancy type hybrid ARQ for finite-length packets in AWGN channel,” in *Proc. IEEE GLOBECOM*, Dec. 2013.